

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-120647

(43)公開日 平成9年(1997)5月6日

(51)Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 1 B 20/12	1 0 2	9295-5D	G 1 1 B 20/12	1 0 2
7/00		9464-5D	7/00	Q
20/10	3 0 1	7736-5D	20/10	3 0 1 Z

審査請求 未請求 請求項の数13 F D (全 18 頁)

(21)出願番号 特願平8-65281

(22)出願日 平成8年(1996)2月27日

(31)優先権主張番号 特願平7-236156

(32)優先日 平7(1995)8月22日

(33)優先権主張国 日本(J P)

(71)出願人 000004329

日本ビクター株式会社

神奈川県横浜市神奈川区守屋町3丁目12番地

(72)発明者 植野 昭治

神奈川県横浜市神奈川区守屋町3丁目12番地 日本ビクター株式会社内

(72)発明者 淵上 徳彦

神奈川県横浜市神奈川区守屋町3丁目12番地 日本ビクター株式会社内

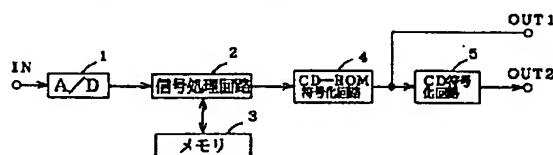
(74)代理人 弁理士 二瓶 正敬

(54)【発明の名称】 オーディオ信号圧縮記録装置及びオーディオ信号圧縮装置並びに光記録媒体

(57)【要約】

【課題】 CD-DA用のデータ圧縮より更に高い圧縮率により、量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化して得たデータをCD-ROMに記録することができるオーディオ信号圧縮記録装置及びオーディオ信号圧縮装置並びに光記録媒体を提供する。

【解決手段】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する手段1、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮する手段2、3、圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域あるいはCD-ROM規格のモード2のユーザデータ領域に配するようフォーマットする手段4を設けた。



## 【特許請求の範囲】

【請求項1】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、

前記データ圧縮手段で圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットするフォーマット手段と、

前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録する手段とを、

有するオーディオ信号圧縮記録装置。

【請求項2】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、

前記データ圧縮手段で圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットするフォーマット手段とを、

有するオーディオ信号圧縮装置。

【請求項3】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体。

【請求項4】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、

前記データ圧縮手段で圧縮されたデータをCD-ROMXA規格のモード2、フォーム2、あるいはCD-ROM規格のモード2のユーザデータ領域に配するようフォーマットするフォーマット手段と、

前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録する手段とを、

有するオーディオ信号圧縮記録装置。

【請求項5】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎

に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、

前記データ圧縮手段で圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域に配するようフォーマットするフォーマット手段とを、

有するオーディオ信号圧縮装置。

【請求項6】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをCD-ROMXA規格のモード2、フォーム2、あるいはCD-ROM規格のモード2のユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体。

【請求項7】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、

前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、

前記データ圧縮手段で圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットするフォーマット手段と、

前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録する手段とを、

有するオーディオ信号圧縮記録装置。

【請求項8】 前記オーディオ信号が2チャンネル信号であり、前記データ圧縮手段が1チャンネルあたり2<sup>m</sup>個(mは正の整数)又はそれ以上の量子化データ毎に前記直交変換及びハフマン符号を適用してデータ量を圧縮するよう構成されている請求項1、4、7のいずれか1つに記載のオーディオ信号圧縮記録装置。

【請求項9】 前記データ圧縮手段が前記オーディオ信号を所定の区間長ごとにフレーム化する手段と、

前記フレーム内の信号を可逆方式で符号化するのに必要な符号量を算出し、フレームで使用可能な符号量と比較する符号量制御手段と、フレーム内の信号を聴覚心理モデルで分析する聴覚心理分析手段と、

フレーム符号量が使用可能符号量以下の場合にはフレーム内の信号を可逆方式で量子化し、フレーム符号量が使用可能符号量を超える場合にはフレーム内の信号を前記聴覚心理分析手段の出力に基づいて非可逆方式で量子化する非可逆量子化手段とを、

有する請求項1、4、7、8のいずれか1つに記載のオーディオ信号圧縮記録装置。

【請求項10】 前記データ圧縮手段が前記オーディオ信号を所定の区間長ごとにフレーム化するフレーム化手

段と、  
符号化対象の全区間の可逆方式による目標符号量と実符号量の差を算出し、各区間毎の符号量の過不足量に応じた補正値を算出する符号量補正値算出手段と、  
フレーム内の信号を聴覚心理モデルで分析する聴覚心理分析手段と、  
全区間の平均符号量が目標符号量になるように前記符号量補正値に基づいて各区間の信号を可逆方式で量子化するか又は前記聴覚心理分析手段の出力に基づいて非可逆方式で量子化する非可逆量子化手段とを、  
有する請求項1、4、7、8のいずれか1つに記載のオーディオ信号圧縮記録装置。

【請求項11】 前記データ圧縮手段が、前記量子化された信号を直交変換する直交変換手段と、前記直交変換手段により直交変換されたデータを複数のバンドに分割するバンド分割手段と、前記バンド分割手段の出力信号に 대응可能なハフマン符号化回路と、前記バンド分割手段によりバンド分割されたデータをバンド毎に前記ハフマン符号化回路に供給する選択手段とを有する請求項1、4、7乃至10のいずれか1つに記載のオーディオ信号圧縮記録装置。

【請求項12】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、  
前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、  
前記データ圧縮手段で圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットするフォーマット手段とを、  
有するオーディオ信号圧縮装置。

【請求項13】 オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、オーディオ信号を高効率符号化して光記録媒体に記録するオーディオ信号圧縮記録装置及びそのためのオーディオ信号圧縮装置、さらに光記録媒体に関する。

【0002】

【従来の技術】CD（コンパクトディスク）は1982年に登場して十数年が経過し、現在では様々な展開によりデジタルストレージメディアとして定着している。オーディオメディアの用途を考えると、サンプリング周

波数 $f_s = 44.1\text{kHz}$ 、量子化ビット数=16ビットのこのメディアは完全に成熟期に入っている。また、この数年のスタジオ製作サイドでは、量子化ビット数の20ビット化、24ビット化や $f_s = 88.2\text{kHz}$ 化、96kHz化等のハイサンプリング化が進んでおり、より高音質のマスタを基にしてCDを作成する動きが出てきている。さらに、DVDと呼ばれる高密度ディスクがコンピュータなどのデータ用のデジタルディスクとして利用されようとしている。なお、デジタルディスクとはCD、CD-ROM、DVDなどオーディオやビデオ信号がデジタル信号として記録された光ディスクをいうものとする。

【0003】

【発明が解決しようとする課題】通常の音楽（オーディオ）用のCD（以下CD-DAという）はサンプリング周波数 $f_s = 44.1\text{kHz}$ 、量子化ビット数=16ビットで2チャンネルのオーディオ信号を記録することができるが、これまでのCD-DAの規格では同一データ量をCD-ROMのフォーマットで記録することができなかった。これは、CD-ROMのフォーマットには同期信号（SYNC）やアドレスやモードを含むヘッダがあるため、オーディオ信号を記録するための記録容量がCD-DAより少ないためである。そのため、音質を向上させるためにハイサンプリング化を行おうとしても、現在のCD-DAを作成するためのデータ処理方式では、データ量が多くなり、CD-ROMに記録することはできなかった。一方、パソコンやその周辺機器の発達と急速な普及により、CD-ROMドライブを介して、音楽などを高音質で楽しみたいという要望がある。

【0004】さらに、DVDと呼ばれるデジタルディスクでは音声がりニアPCMにより圧縮されずに記録されているため、よりハイファイ性の高い記録のためにはデータ量を要し、記録時間が短くなる。このディスクのためには、直交変換及び／又はハフマン符号によりデータ処理してデータ量を削減するための圧縮を行って、DVDのフォーマットで記録する記録装置並びにかかる方式で記録された光ディスクが考えられる。

【0005】したがって、本発明は現在のCD-DA用のデータ処理に比較して高い圧縮率により、量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数で量子化して得たデータをCD-ROMに記録することができるオーディオ信号圧縮記録装置及びオーディオ信号圧縮装置並びに光記録媒体を提供することを目的とする。

【0006】また、本発明は、CDバリエーション（サイズ・変調方式）の範囲内でデータフォーマットが一般に異なると見られているDVDオーディオに記録することができるオーディオ信号圧縮記録装置及びオーディオ信号圧縮装置並びに光記録媒体を提供することを目的とする。

## 【0007】

【課題を解決するための手段】本発明は上記目的を達成するために、オーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を削減・圧縮し、圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域又はCD-ROM規格のモード2のユーザデータ領域に配するようフォーマットし、フォーマットされたデータを記録媒体に記録するようにしている。また、本発明によれば光記録媒体としての高密度記録の可能な新規格のDVDが提供され、DVDのユーザデータ領域に上記の直交変換及びハフマン符号を適用してデータ量を削減・圧縮し、圧縮されたデータを記録するようにしている。

【0008】すなわち、本発明によれば、オーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットするフォーマット手段と、前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録する手段とを、有するオーディオ信号圧縮記録装置が提供される。

【0009】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域に配するようフォーマットするフォーマット手段と、前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録する手段とを、有するオーディオ信号圧縮記録装置が提供される。

【0010】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットするフォーマット手段と、前記フォーマット手段でフォーマットされたデータをCDフォーマットとして記録媒体に記録

する手段とを、有するオーディオ信号圧縮記録装置が提供される。

【0011】なお、本発明は上記のようにオーディオ信号圧縮記録装置として捉えられるが、さらに、再生専用のディスクの製造のためには、CDフォーマットとして記録する工程はディスク製造工場側のタスクとなる。したがって、本発明はオーディオ信号圧縮装置としても捉えることができる。

【0012】すなわち、オーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットするフォーマット手段とを、有するオーディオ信号圧縮装置が提供される。

【0013】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域に配するようフォーマットするフォーマット手段とを、有するオーディオ信号圧縮装置が提供される。

【0014】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段と、前記量子化手段で量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮するデータ圧縮手段と、前記データ圧縮手段で圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットするフォーマット手段とを、有するオーディオ信号圧縮装置が提供される。

【0015】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをデジタルディスクのユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体が提供される。

【0016】また、本発明によれば、オーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータ

をCD-ROMXA規格のモード2、フォーム2のユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体が提供される。

【0017】また、本発明によればオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをDVDのユーザデータ領域に配するようフォーマットし、フォーマットされたデータをCDフォーマットとして記録した光記録媒体が提供される。

【0018】

【発明の実施の形態】本発明のオーディオ信号圧縮記録装置及びオーディオ信号圧縮装置並びに光記録媒体の実施の形態を好ましい実施例によって説明する。図1は本発明のオーディオ信号圧縮記録装置の好ましい実施例を示すブロック図である。入力端子INには例えば音楽信号などのアナログ信号が供給され、出力端子OUTは図示省略のCD原盤作成機、すなわちマスタリング装置に必要に応じてプリマスタリング装置を介して接続される。マスタリング装置自体は従来のものと本質的に変わらないので、ここでは説明を省略する。

【0019】図1の装置は入力端子INに接続されたA/D変換器1と、その出力に接続された信号処理回路2と、信号処理回路2に接続されたメモリ3と、信号処理回路2の出力に接続されたCD-ROM符号化回路4と、CD-ROM符号化回路4の出力に接続されたCD符号化回路5を有している。CD-ROM符号化回路4の出力は第1出力端子OUT1に接続され、CD符号化回路5の出力は第2出力端子OUT2に接続されている。なお、後述するように、CD符号化回路5は不要な場合がある。

【0020】A/D変換器1はオーディオ信号を量子化ビット数20ビット又はそれ以上、標本化周波数44.1kHz又はそれ以上の周波数で量子化する量子化手段として動作する。標本化周波数は実施例により44.1kHz（DVDの場合は48kHz）又は88.2kHz（DVDの場合は96kHz）のいずれかになっているが、44.1kHz以上の適当な値とすることができる。音楽信号を対象とする場合は、通常左右の2チャンネルであるが、サラウンドその他の必要に応じて4チャンネルや6チャンネルなどとする事ができる。ここでは2チャンネルである場合について説明する。A/D変

換器1で得られた量子化データは1チャンネルあたり2<sup>m</sup>個（mは正の整数）を単位として、信号処理回路2を介してメモリ3に書込まれる。その後、信号処理回路2がこの2<sup>m</sup>個のデータの処理を開始する。

【0021】図2は信号処理回路2の一例を示すブロック図である。2<sup>m</sup>個のデータは直交変換回路10にて直交変換が施され、周波数スペクトルが得られる。この周波数スペクトルをバンド分割のための複数のフィルタ6a, 6b, 6c, . . . 6nを有するフィルタバンク6と選択手段としてのスイッチ回路7を介して正規化部・量子化部11に与え、バンド毎にまとめて正規化・量子化する。ここで正規化レベル（ビット数）を補助情報、スペクトルデータを主情報としてデータフレームとする。このデータフレームをハフマン符号化回路8に与えて、ハフマン符号化処理を行い、データ量を削減・圧縮するとともに、コードブックのインデックスを補助情報、処理データを主情報として、新たなデータフレームを作成し、これを順次メモリ3に書き込む。次にメモリ3からこの新たなデータフレームを読み出し、アロケーション回路9を介して図1のCD-ROM符号化回路4へ出力する。

【0022】CD-ROM符号化回路4では、図3によって後述する所定のフォーマットとなるように、各セクタに同期信号（SYNC）やヘッダ、サブヘッダなどを付加し、各セクタのユーザデータ領域に信号処理回路2から与えられる圧縮オーディオデータを配して出力する。CD-ROM符号化回路4の出力データは第1出力端子OUT1を介して出力され、例えば磁気テープに記録されて、再生専用のCDを製造するためのプリマスタリング装置やマスタリング装置に供給される。一方、CD-ROM符号化回路4の出力データは、書込み可能な、いわゆるライトワンスタイプのCDの場合は、CD符号化回路5に与えられ、CDフォーマット化され、第2出力端子OUT2を介して図示省略の記録ヘッドにより記録される。

【0023】次に図3と共に本発明のいくつかの態様について説明する。図3はCDの種々のフォーマットをセクタ単位で示したもので、第1段には通常の音楽用CDである、CD-DAを示し、以下第2段から第6段まで各種CD-ROMを示している。本発明の実施例としては次の6つの態様がある。なお、DVDを示す図17については後述する。

【0024】

【表1】

- |     |                               |
|-----|-------------------------------|
| (1) | CD-ROM XA モード2、フォーム2 （図3の6段目） |
|     | 標本化周波数 : 44.1kHz              |
|     | 量子化ビット数 : 20ビット               |
| (2) | CD-ROM XA モード2、フォーム2 （図3の6段目） |
|     | 標本化周波数 : 88.2kHz              |
|     | 量子化ビット数 : 20ビット               |

- (3) CD-ROM モード2 (図3の4段目)  
 標本化周波数 : 44.1 kHz  
 量子化ビット数 : 20ビット
- (4) CD-ROM モード2 (図3の4段目)  
 標本化周波数 : 88.2 kHz  
 量子化ビット数 : 20ビット
- (5) DVD (図17)  
 標本化周波数 : 48 kHz  
 量子化ビット数 : 20ビット～24ビット
- (6) DVD (図17)  
 標本化周波数 : 96 kHz  
 量子化ビット数 : 20ビット～24ビット

【0025】CD-ROM XA モード2、フォーム2ではユーザデータは2324バイトである。また、CD-ROM モード2では、ユーザデータは2336バイトである。これらの規格では、比較的ユーザデータのデータ量、すなわちバイト数が多いので、1枚のディスクに記録収納可能なデータ量が多く、有利である。

【0026】また、上記(1)、(2)のCD-ROM XA モード2、フォーム2を用いた場合は、独自の割当てのサブヘッダを規定することができる。サブヘッダの内容を表2に示す。

【0027】

【表2】

サブヘッダ

バイト No.	バイト値
16	ファイル No.
17	チャンネル No.
18	サブモード
19	コーディング情報
20	ファイル No.
21	チャンネル No.
22	サブモード
23	コーディング情報

【0028】上記サブヘッダ中、サブモードバイトのビ

ット5～2をこの符号化IDに用いることで、サブヘッダを見ながら、このフォーマットのデコードを行うことができる。以下の表3と表4に、サブヘッダ中のサブモードと、コーディング情報の内容を示す。サブヘッダにはフォーマット時の条件を記録することができるが、その手法として2つの方法がある。その一つはそのセクタのフォーマット条件を入れる方法であり、他の方法はフォーマット条件を複数のセクタに分けて記録する方法であり、この場合これら複数のセクタの情報を集合して解読可能となる。

【0029】

【表3】

サブモード バイト

ビット No.	ビット名
7	エンド オブ ファイル (EOF)
6	リアルタイム セクタ (RT)
5	フォーム (F)
4	トリガ (T)
3	データ (D)
2	オーディオ (A)
1	ビデオ (V)
0	エンド オブ レコード (EOR)

【0030】

【表4】

## コーディング情報バイト

ビット セクタ	7	6 5 4 3 2						1	0	
オーディオ	0	Emph	ADPCM レベル						モード	
		0=Off	0000=レベル B						00=モノ	
		1=On	0001=レベル C						01=ステレオ	
ビデオ	0	X	X	X	X	X	X	X	CDI	
	1	0	0	0	0	0	0	0	0	ASM
			レゾリューション			コーディング			EVM	
			000=320X200 001=640X480			001=CLUT1 010=CLUT2 011=CLUT4 100=CLUT8				
			1	リザーブ						
データ	リザーブ									

Emph: エンファシス ビット  
 レベルB: 4ビット37.8KHzサンプリングレート  
 レベルC: 4ビット18.9KHzサンプリングレート  
 CDI: CD-Iビデオモード  
 ASM: アプリケーション特定ビデオモード  
 EVM: 拡張ビデオモード

【0031】上記4つの態様中、標準化周波数が88.2kHzである、(2)と(4)では、2ブロックで1フレームを構成することとなる。したがって、44.1kHzの場合と比較して、記録できる時間は半分となる。

【0032】上記実施例は、信号処理回路2が可逆圧縮方式である場合について説明したが、本発明者らが先に開発したいわゆる準可逆符号化方式のものを適用することにより、更にデータ量を圧縮することができる。以下にこの方式について説明する。

【0033】図4は図1の信号処理回路に本発明者らが開発した音声の準可逆符号化装置を適用する場合の例を示すブロック図、図5は図4における聴覚心理分析と符号量調整処理を説明するためのフローチャート、図6は図4の準可逆符号化装置と従来例における符号量不足時の再量子化ノイズレベルの比較例を示す説明図、図7は図4の準可逆符号化装置と従来例における聴感上の音質比較例を示す説明図である。

【0034】図4に示す装置ではまず、従来の周波数領域処理のエンコーダと同様に、バッファ21が後段の窓掛け・直交変換部22が直交変換する際に必要なフレーム分のPCM信号をバッファリングし、窓掛け・直交変換部22はこのフレームデータに窓掛け（一般にはハニング窓等の窓掛け）し、MDCT（変形離散コサイン変換）等により直交変換し、この直交変換係数を複数のバンドに分割する。正規化部23はこのバンド毎の正規化係数（スケールファクタ）を決定し、バンド内の直交変換係数を正規化する。量子化・符号化部24はこの正規化後の係数を可逆に必要な精度で量子化し、必要であればエントロピー符号化する。

【0035】そして、図4の例では、聴覚心理分析部25と符号量制御部26及び量子化・符号化部24が図5に示すような処理を行う。図5において、まず、量子化

・符号化部24により正規化された係数の1回目の量子化ビット数 (Bit[i]) を決定し、符号量を見積もって総符号量 (Total bit) を算出する (ステップS1)。次いでそのフレームの使用可能符号量 (Avail bit) を確認又は算出し (ステップS2)、次いで総符号量 (Total bit) と使用可能符号量 (Avail bit) を比較することにより符号量が不足するか否かをチェックする (ステップS3)。

【0036】そして、符号量が不足する場合 (Total bit > Avail bit) には、まず、聴覚心理モデルのマスキング効果と最小可聴限特性を考慮してバンドパワー  $p[i]$  ( $=$  正規化値<sup>2</sup>  $=$  scale[i]<sup>2</sup>) からマスキングカーブ  $m[i]$  を算出する (ステップS4)。この場合、マスキングカーブ  $m[i]$  は基準カーブ  $curve[i]$  とバンドパワー  $p[i]$  を畳み込み演算することにより得られる。

【0037】次いで最小可聴限とマスキングカーブから各バンドの標準ノイズレベル  $N[i]$  を算出し (ステップS5)、次いで標準ノイズレベル  $N[i]$  が高いバンドから1ビットずつビット削減を行うことにより不足符号量を各バンドに振り分ける。但し、バンド  $i$  において1ビット削減を行う毎に  $N[i]$  から6.0を減算し、ビット削減が標準ノイズレベル  $N[i]$  と相似形になるようにする (ステップS6)。そして、このように各バンド毎に最終的に決定された量子化ビット数で、量子化・符号化部24で再量子化及び符号化する (ステップS7)。

【0038】また、ステップS3において符号量が不足しない場合には、余剰ビットを各バンドに割り当て又はパディングし (ステップS8)、その量子化ビット数で、量子化・符号化部24で再量子化及び符号化する (ステップS7)。フォーマット出力部26は一般に、正規化係数 (場合によっては量子化ビット数) と、符号量制御部26の符号量制御情報と、それにヘッダ等の補助情報を付加してフォーマット化 (ビットストリーム

化)して伝送する。

【0039】図6は図4の例と、従来のエンコーダにおいて符号量不足時の再量子化ノイズレベルの設定例を比較した場合を示している。上記例によれば、再量子化ノイズ聴覚心理モデルに応じてシェーピングされており、ノイズ量が同じであっても聴感上ではノイズレベルが下がった場合と同等の効果を得ることができる。したがって、聴感上の音質劣化を最小限にして準可逆的に符号化することができる。

【0040】図7は従来例で非可逆符号化を行った場合と、上記例の場合の音質の比較例を示し、図7(a)はフレームの一部が非可逆となる場合、図7(b)はフレームの大部分が非可逆となる場合を示す。図のように非可逆となる区間において太線で示す本発明の方が細線で示す従来例より音質を改善することができ、したがって、符号化全体として安定した音質を得ることができる。また、本発明によれば、非可逆符号化を行った場合の音質を十分確保することができるので、各フレームの使用可能符号量が一定の「固定伝送レート」で伝送することができ、したがって、非可逆フレームが大幅に増加しても音質上の問題は発生しない。この結果、オーサリングや再生装置側の符号量制御に関わる処理を大幅に簡略化することができる。

【0041】次に、信号処理回路の他の例について図8乃至図14に沿って説明する。図8は図4同様、本発明者らが開発した音声の準可逆符号化装置の他の例を示すブロック図、図9は図8における符号量補正値を算出する処理を説明するためのフローチャート、図10は符号量偏差と符号量補正値の関係を示すグラフ、図11～図13は符号量補正前と補正後の符号量偏差ヒストグラムを示す説明図、図14は図8における聴覚心理分析と符号量調整処理を説明するためのフローチャート、図15は図8の準可逆符号化装置と従来例における符号量過剰時の再量子化ノイズレベルの比較例を示す説明図、図16は図8の準可逆符号化装置と従来例における聴感上の音質比較例を示す説明図である。

【0042】図8に示す装置では、まず、従来の周波数領域処理のエンコーダと同様に、バッファ21が後段の窓掛け・直交変換部22が直交変換する際に必要なフレーム分のPCM信号をバッファリングし、窓掛け・直交変換部22はこのフレームデータに窓掛け(一般にはハニング窓等の窓掛け)し、MDCCT(変形離散コサイン

変換)等により直交変換し、この直交変換係数を複数のバンドに分割する。正規化部23はこのバンド毎の正規化係数(スケールファクタ)を決定し、バンド内の直交変換係数を正規化する。量子化・符号化部24はこの正規化後の係数を可逆に必要な精度で量子化し、この場合にも必要であればエントロピー符号化する。但し、図11に示す時間領域処理の場合よりエントロピー符号化の効果は一般に少ない。

【0043】そして、この例では、聴覚心理分析部25と符号量制御部26及び量子化・符号化部24が区間毎の符号量補正値Adjに基づいて以下のような処理を行う。まず、本発明では、オーディオメディアを制作する場合に、1曲(例えば4～6分)又は全曲(例えば40～74分)等の長時間平均で符号量が目標値になるように制御する方法であり、エンコード処理は2パスで行う。具体的には、

(a)可逆符号化を仮定した1回目のエンコード処理を行う。但し、各区間の使用符号量が得られればよく、実際に量子化・符号化を行う必要はない。

(b)図9に示すように各区間の使用符号量と目標符号量の差から各区間の符号量補正値Adjを算出する。

(c)2回目のエンコード処理を行う。この場合、可逆符号化を仮定したビット割り当てを補正符号量と聴覚心理モデルにより変更して量子化・符号化を行い、また、ビット割り当て変更の情報を補助情報としてデコードに伝送する。

【0044】次に、図9を参照して上記(b)における符号量補正値Adjを算出する処理について説明する。

①まず、対象区間の使用符号量を入力して平均符号量 $T_m$ を算出し、目標符号量 $T_d$ との差を評価する(ステップS11、S12)。

②次いで、符号量過剰な場合(平均符号量 $T_m >$  目標符号量 $T_d$ )には、各区間の使用符号量と目標符号量との偏差 $\Delta[\text{bit}]$ (但し、過剰な場合に正)を算出し、この偏差 $\Delta[\text{bit}]$ を適当なステップ幅 $\text{step}[\text{bit}]$ で量子化し、ヒストグラムを作成する(ステップS12→S13)。

③次いで、ヒストグラムの偏差が負の領域の偏差総量 $S_m$ と、正の領域の偏差総量 $S_p$ を以下のように算出する(ステップS14)。

【0045】

【数1】

$$S_m = \sum_{i=\min}^{-1} \text{histogram}[i] * |i| \quad [\text{step}]$$

但し、 $i$ はヒストグラムのインデックス  
 $\min, \max$ はインデックスの下限、上限

$$S_p = \sum_{i=1}^{\max} \text{histogram}[i] * i \quad [\text{step}]$$

【0046】④次いで、負の領域の偏差総量 $S_m$ の比率 $S_m / (S_m + S_p)$ が予め定めた値Bound(例えば0.33等)より大きい場合には、以下のように各区間

毎の符号量補正値Adjを求める(ステップS15→S16)。

【0047】

【数2】

if  $\Delta \leq 0$ 

Adj = 0 [bit]

else

Adj = - { (Sp - Sm) / Sp } \*  $\Delta$ [bit]

【0048】④' 他方、比率  $Sm / (Sm + Sp)$  が値 Bound より小さい場合には、比率  $Sm / (Sm + Sp)$  が値 Bound より大きくなるようにヒストグラムのオフセット値 Off を決定し（ステップ S15→S17）、以下のように各区間毎の符号量補正值 Adj を求める（ステップ S18）。

【0049】

【数3】

if  $\Delta \leq Off * step$ 

Adj = -Off \* step [bit]

else

Adj = -Off \* step

- { (Sp - Sm) / Sp }

\* ( $\Delta - Off * step$ ) [bit]

ここで、この手法を用いる理由は、ヒストグラムが極端に「過剰」側に偏っている場合には、ある程度全フレームにオフセットを掛けて補正する必要があるからである。

【0050】②' また、ステップ S12 において平均符号量  $T_m > 目標符号量 T_d$  でない場合には、平均符号量  $T_m$  と目標符号量  $T_d$  に基づいて以下のように各区間で一定の符号量補正值 Adj を求める（ステップ S19）。

Adj = (Td -  $T_m$ ) [bit]

【0051】図10は符号量偏差  $\Delta$ [bit] と符号量補正值 Adj の関係を示し、偏差  $\Delta$ [bit] が正であって大きい程、補正值 Adj も増大する。また、図11～図13は符号量補正前（実線）と補正後（破線）のヒストグラムを示し、横軸がサンプル当たりの偏差 ( $\Delta$  / 区間当たりのサンプル数) を、また、縦軸が度数を示す。詳しくは図11は上記④のように補正值 Adj を求めた場合、また、図12、図13はそれぞれ上記④'、②'のように補正值 Adj を求めた場合を示している。

【0052】次に、図14を参照して聴覚心理分析と符号量調整処理を説明する。図14において、まず、量子化・符号化部24により正規化された係数の1回目（可逆方式）の量子化ビット数 ( $Bit[i]$ ) を決定し、符号量を見積もって総符号量 (Total bit) を算出する（ステップ S21）。次いでそのフレームの符号量補正值 Adj を読み込み（ステップ S22）、補正值 Adj が負 ( $Adj < 0$ ) か否かをチェックする（ステップ S23）。

【0053】そして、補正值 Adj が負の場合（符号量削減）には、まず、聴覚心理モデルのマスクング効果と最小可聴限特性を考慮してバンドパワー  $p[i]$  (=正規化値<sup>2</sup> =  $scale[i]^2$ ) からマスクングカーブ  $m[i]$  を算出する（ステップ S4）。この場合、マスクングカーブ  $m$

[i] は基準カーブ  $curve[i]$  とバンドパワー  $p[i]$  を畳み込み演算することにより得られる。

【0054】次いで最小可聴限とマスクングカーブから各バンドの標準ノイズレベル  $N[i]$  を算出し（ステップ S5）、次いで標準ノイズレベル  $N[i]$  が高いバンドから1ビットずつビット削減を行うことにより符号量補正值を各バンドに振り分ける。但し、バンド i において1ビット削減を行う毎に  $N[i]$  から6.0を減算し、ビット削減が標準ノイズレベル  $N[i]$  と相似形になるようにする（ステップ S6）。そして、このように各バンド毎に最終的に決定された量子化ビット数で、量子化・符号化部24で再量子化及び符号化する（ステップ S7）。

【0055】また、ステップ S23 において補正值 Adj が負でない場合（符号量増加）には、余剰ビットを各バンドに割り当て又はパディングし（ステップ S8）、その量子化ビット数で、量子化・符号化部24で再量子化及び符号化する（ステップ S7）。フォーマット出力部26は一般に、正規化係数（場合によっては量子化ビット数）と、符号量制御部26の符号量制御情報と、それにヘッダ等の補助情報を付加してフォーマット化（ビットストリーム化）して伝送する。

【0056】したがって、上記例によれば、算術エントロピーが大きく、聴感エントロピーが小さい区間ほど、より多くの符号量補正（削減）を受けることになり、聴感に対応した符号量配分を行うことができる。また、図15は図8の装置と、従来のエンコーダにおいて符号量過剰時の再量子化ノイズレベルの設定例を比較した場合を示し、この例によれば、非可逆符号化されるフレームにおいても再量子化ノイズ聴覚心理モデルに応じてシェーピングされており、ノイズ量が同じであっても聴感上ではノイズレベルが下がった場合と同等の効果を得ることができる。したがって、聴感上の音質劣化を最小限にして準可逆的に符号化することができる。

【0057】図16は従来例において非可逆符号化を行った場合と、上記図8の例の場合の音質の比較例を示し、図16(a)はフレームの一部が非可逆となる場合、図16(b)はフレームの大部分が非可逆となる場合を示す。図のように非可逆となる区間において太線で示す本発明の方が細線で示す従来例より音質を改善することができる。したがって、符号化全体として安定した音質を得ることができる。

【0058】次に本発明によるデータ圧縮率がどの程度であるかについて検討する。表5は音楽の3つのジャンル別に、圧縮効果を実測あるいは予測した結果を示したものである。なお、表中の1段目の4桁の数値の上2桁はビット数を、下2桁は標準化周波数を示している。各ジャンルにおいて、5乃至10曲を選定して調査した。

【0059】

【表5】

	1644	2044	2088高域予想	2088予想
ポップス	70~77~83	76~82~86	33~40~50	55~61~68
ジャズフュージョン	58~65~77	66~72~82	25~33~50	48~53~66
クラシック	40~50~60	52~60~68	20~28~38	36~43~53

圧縮前のデータ量を100とした圧縮後のデータ量

【0060】表5中の数字は基のデータ量を100としたときの圧縮後のデータ量を示している。この表から分るように、例えば、2044をクラシック音楽に適用すると、平均で40%、最大で48%の圧縮が可能であり、さらに2088をクラシック音楽に適用すると、平均で57%、最大で64%の圧縮が可能であることが分る。ポップスやジャズ・フュージョンではクラシック程の圧縮はできないが、平均的に2044では18%から28%、2088では39%から47%の圧縮率が得られる。なお、表中、2088高域予想とは16kHz~20kHz又は13kHz~20kHzのスペクトルのスロープの傾向から20kHz以上の高域について予測したものである。

【0061】表5に示した圧縮の効果は、図4及び図8に示した準可逆符号化装置を用いない場合のものであり、準可逆符号化装置を用いることにより、さらに圧縮率を高くすることができる。

【0062】上記本発明の実施例の4つの態様は、そのいずれかを選択できるように、本発明の圧縮記録装置又は圧縮装置の使用者が手動で図示省略のセレクトボタンなどを操作することにより、切り替えて使用できる構成とすることができる。なお、標準化周波数を44.1kHzより高く設定した場合は、44.1kHzのときの一定線速度より更に速い線速度となるよう、ディスクの回転数を制御する必要がある。標準化周波数を44.1kHzより高く設定した場合は、高域の周波数特性が改善され、高音質化が図られる。図17はDVDのフォーマットを図3と同様にセクタ単位で示すデータ配置模式図である。図17に示されるように、DVDでは通常1バックが2048バイト(1論理セクタ)で構成され、その中のパケット(ユーザデータ)2034バイトが利用できる。なお、図17において、「バックスタート」は同期信号となるSYNCパターンを有し、「SCR」は時間情報であるシステム・クロック・レファレンスであり、「Mux rate」は転送レート(マルチプレクシングレート)であり、「パケット(ユーザデータ)」はパケットヘッダとデータなどからなる。

【0063】

【発明の効果】以上説明したように本発明によれば、オーディオ信号を量子化ビット数20ビット又はそれ以上、標準化周波数44.1kHz又はそれ以上の周波数

で量子化し、量子化された所定量の量子化データ毎に直交変換及びハフマン符号を適用してデータ量を圧縮し、圧縮されたデータをCD-ROMXA規格のモード2、フォーム2のユーザデータ領域、あるいはCD-ROM規格のモード2のユーザデータ領域、あるいはDVDのユーザデータ領域に配するようフォーマットングするようにしているので、音声信号を高圧縮率で圧縮して、CD-ROMやDVDに記録することができ、CD-ROMオーディオやDVDオーディオを実現することができる。

【図面の簡単な説明】

【図1】本発明のオーディオ信号圧縮記録装置の好ましい実施例を示すブロック図である。

【図2】図1中の信号処理回路2の一例を示すブロック図である。

【図3】CDの種々のフォーマットをセクタ単位で示したデータ配置模式図である。

【図4】図1中の信号処理回路2の他の例としての音声の準可逆符号化装置の一例を示すブロック図である。

【図5】図4における聴覚心理分析と符号量調整処理を説明するためのフローチャートである。

【図6】図4の準可逆符号化装置と従来例における符号量不足時の再量子化ノイズレベルの比較例を示す説明図である。

【図7】図4の準可逆符号化装置と従来例における聴感上の音質比較例を示す説明図である。

【図8】図1中の信号処理回路2のさらに他の例としての音声の準可逆符号化装置の一例を示すブロック図である。

【図9】図8における符号量補正値を算出する処理を説明するためのフローチャートである。

【図10】符号量偏差と符号量補正値の関係を示すグラフである。

【図11】符号量補正前と補正後の符号量偏差ヒストグラムを示す説明図である。

【図12】符号量補正前と補正後の符号量偏差ヒストグラムを示す説明図である。

【図13】符号量補正前と補正後の符号量偏差ヒストグラムを示す説明図である。

【図14】図8における聴覚心理分析と符号量調整処理を説明するためのフローチャートである。

【図15】図8の準可逆符号化装置と従来例における符号量不足時の再量子化ノイズレベルの比較例を示す説明図である。

【図16】図8の準可逆符号化装置と従来例における聴感上の音質比較例を示す説明図である。

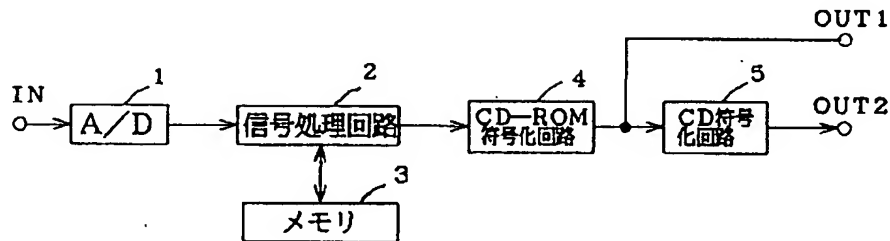
【図17】DVDのフォーマットをセクタ単位で示したデータ配置模式図である。

【符号の説明】

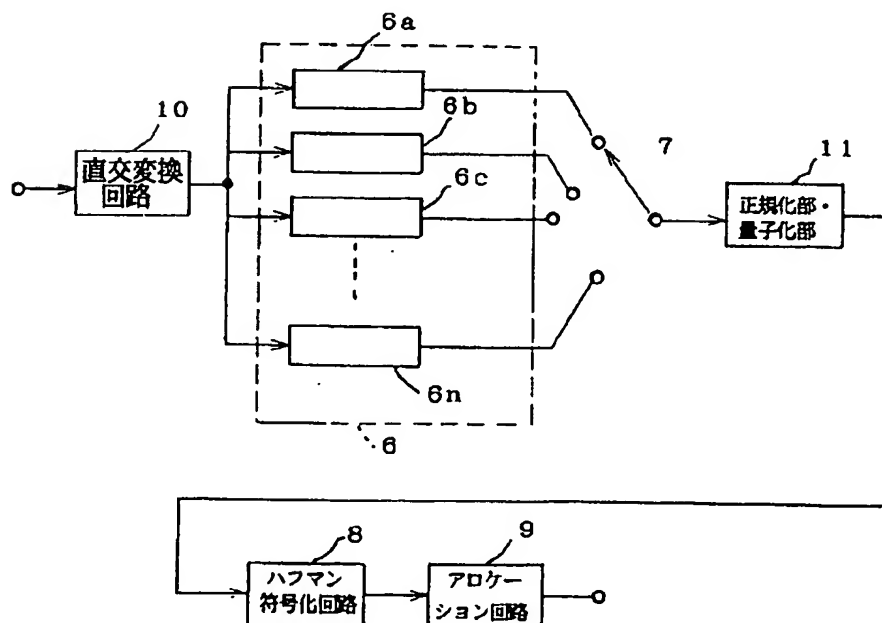
- 1 A/D変換回路（量子化手段）
- 2 信号処理部（メモリ3とともにデータ圧縮手段を構成する）
- 3 メモリ
- 4 CD-ROM符号化回路（フォーマット化手段）
- 5 CD符号化回路

- 6 フィルタバンク（バンド分割手段）
- 7 スイッチ回路（選択手段）
- 8 ハフマン符号化回路
- 9 アロケーション回路
- 10 直交変換回路
- 11 正規化部・量子化部
- 21 バッファ
- 22 窓掛け・直交変換部
- 23 正規化部
- 24 量子化・符号化部
- 25 聴覚心理分析部
- 26 符号量制御部
- 27 フォーマット化出力部
- IN 入力端子
- OUT1、OUT2 出力端子

【図1】



【図2】



【図3】

[CD DA]  
レッドブック (IEC908): CD-DA

2352 バイト									
L	L			L	L			R	R
LSB	MSB	LSB	MSB	LSB	MSB			LSB	MSB
(1)	(1)	(1)	(1)	(1)	(1)			(1)	(1)

[CD-ROM モード0]  
イエローブック (ISO/IEC 10149)

					2352 バイト	
SYNC (12)	ヘッダ (4)				"00" (2336)	
	アドレス (3)					
	分	秒	ミクロ秒	モード		
	(1)	(1)	(1)	(1)		

[CD-ROM モード1]  
イエローブック (ISO/IEC 10149)

2352 バイト									
SYNC (12)	ヘッダ (4)				ユーザデータ (2048)	E D C (4)	00 (8)	ECC	
	アドレス (3)							Pパリティ	Qパリティ
	分 (1)	秒 (1)	ミクロ ク (1)	モ ド ー ド (1)				(172)	(104)

[CD-ROM モード2]  
イエローブック (ISO/IEC 10149)

					2352 バイト	
SYNC (12)	ヘッダ (4)				ユーザデータ (2336)	
	アドレス (3)					
	分	秒	ミリ 秒	モード		
	(1)	(1)	(1)	(1)		

[CD-ROM XA]  
CD-ROM モード2 フォーム1

2352 バイト									
SYNC (12)	ヘッダ (4)				サブ ヘッダ モード (1)	ユーザデータ (2048)	E D C (4)	ECC	
	アドレス (3)			Pパリティ				Qパリティ	
	分	秒	ミクロ	(172)				(104)	
	(1)	(1)	(1)	(1)	(8)				

[CD-ROM XA]  
CD-ROM モード2 フォーム2

2352 バイト									
SYNC (12)	ヘッダ (4)				サ ハ ヘ ッ ダ	ユーザデー タ	E D C (4)		
	アドレス (3)								
	分	秒	ミ ク ロ 秒	モ ー ド					
	(1)	(1)	(1)	(1)					
(2324)									

(a)

音質

時間

本発明

従来

許容逆区間

(b)

音質

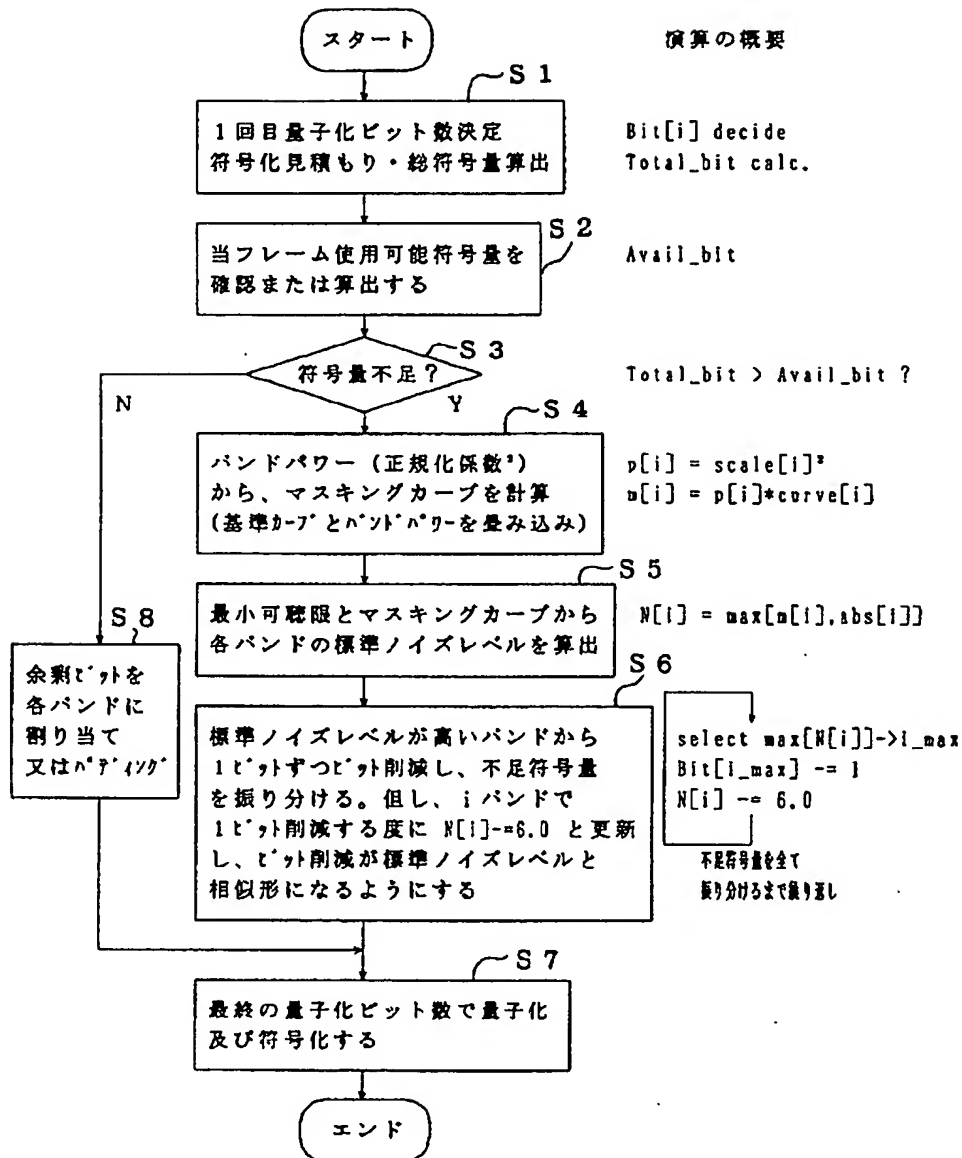
時間

本発明

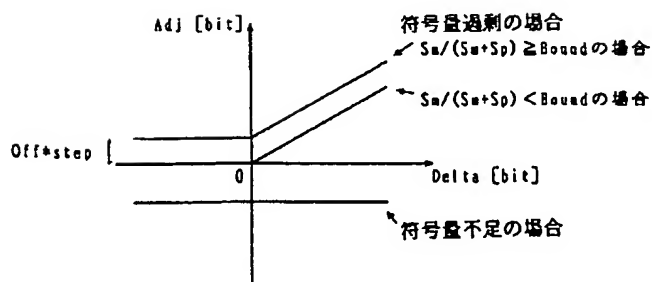
従来

許容逆区間

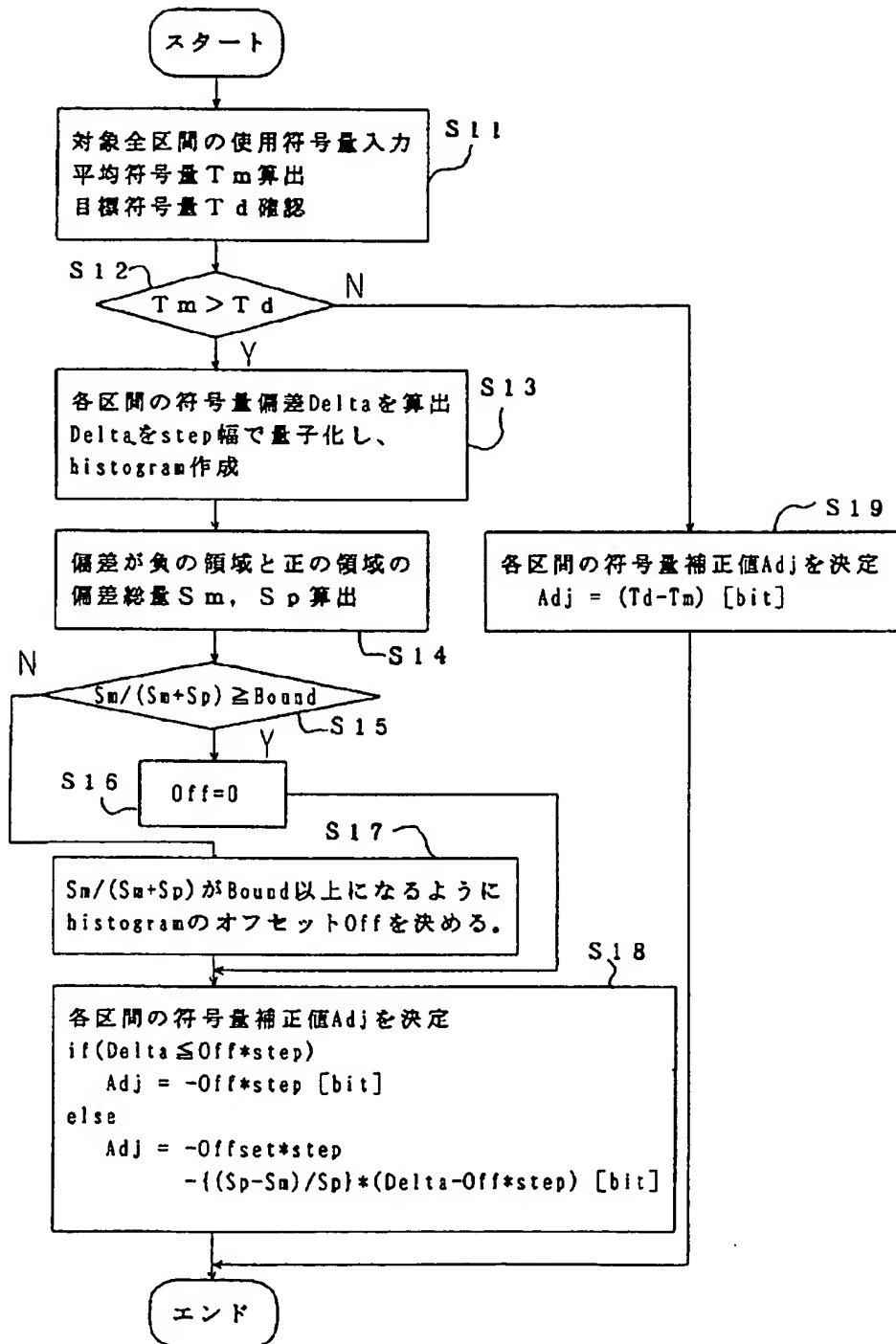
【図5】



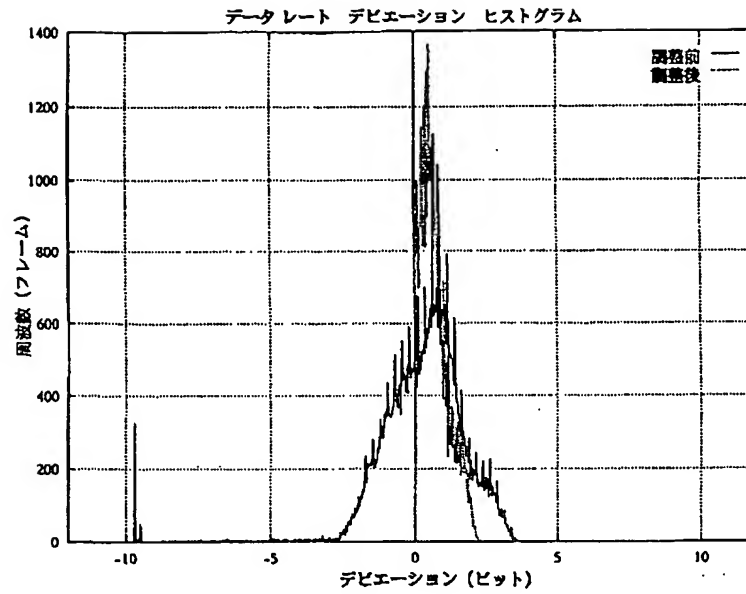
【図10】



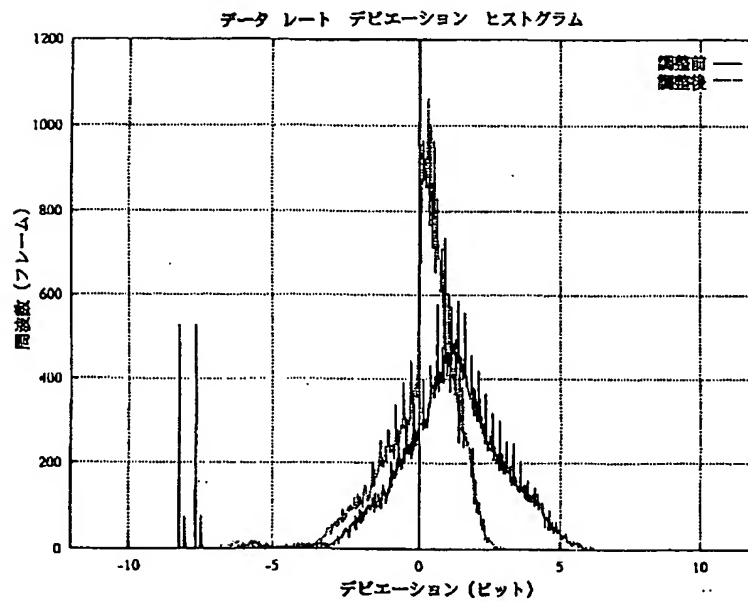
【図9】



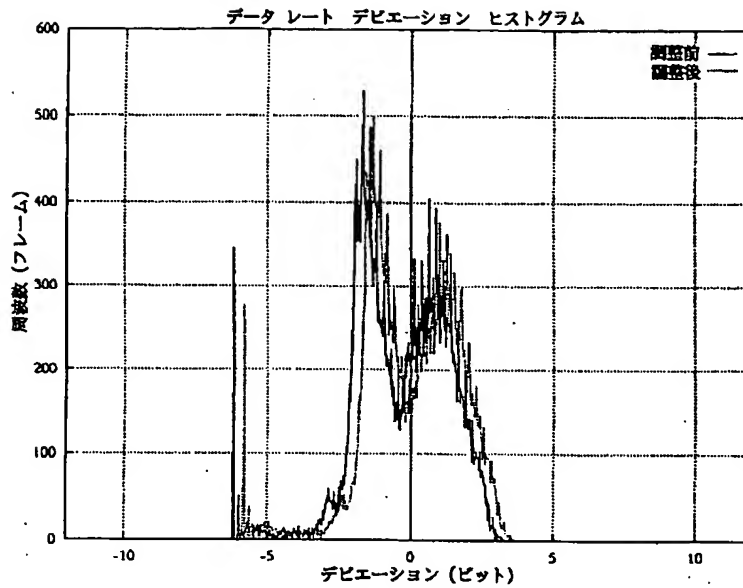
【図11】



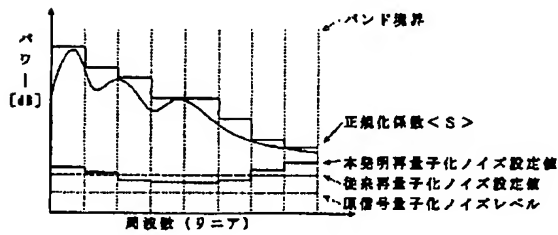
【図12】



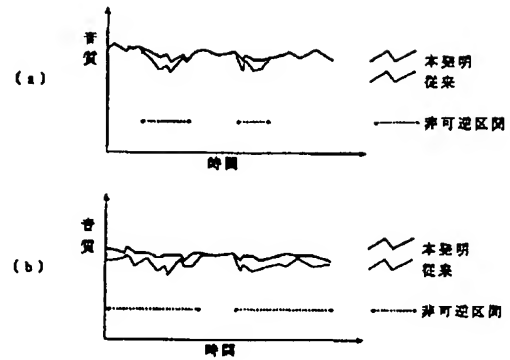
【図13】



【図15】



【図16】



【図17】

[DVD]

2048バイト				
バック スタート	SCR	Mux rate	スタッフィング	パケット (ユーザデータ)
(4)	(8)	(3)	(1)	(2034)

バックヘッダ (14)

【図14】

